AI In Medicine: Considerations for Building Trust

Transparency, Reproducibility, Performance, Clinical Impact

Sameer Antani, PhD

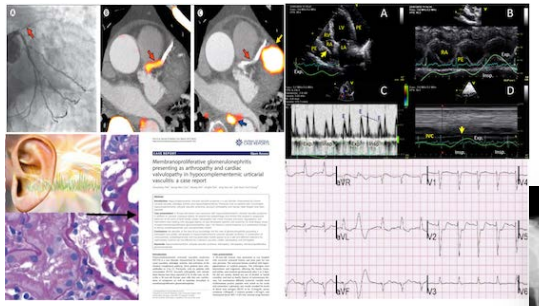
Staff Scientist, Lead Image Analytics and Machine Learning Group

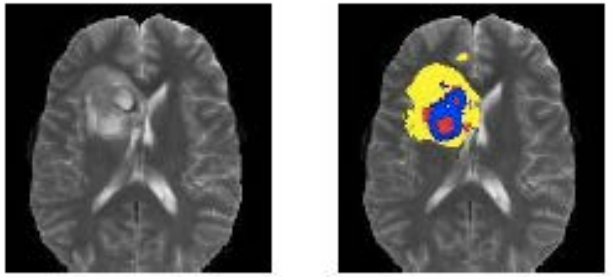
Lister Hill National Center for Biomedical Communications

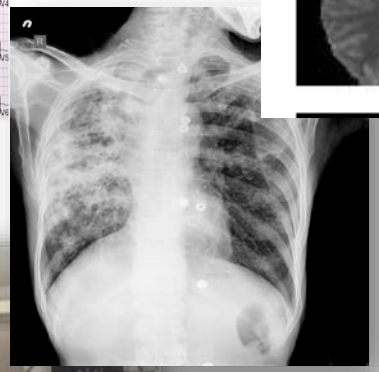
National Library of Medicine, NIH

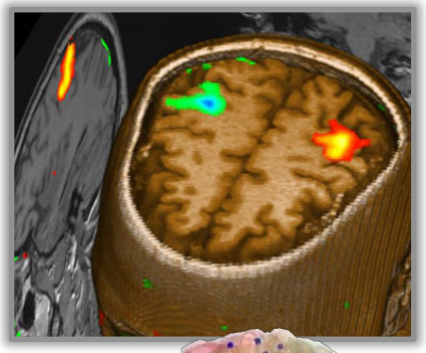
BIOMEDICAL IMAGING ML/AI PROJECTS

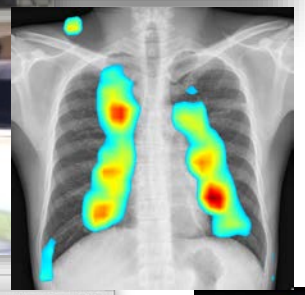
CADx, INFORMATICS, GLOBAL HEALTH, IMAGE DATABASES, QUESTION ANSWERING,

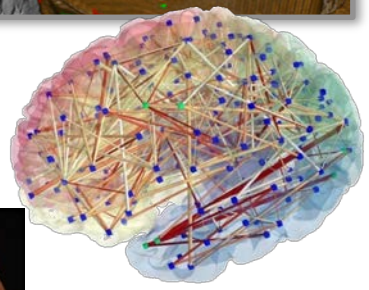


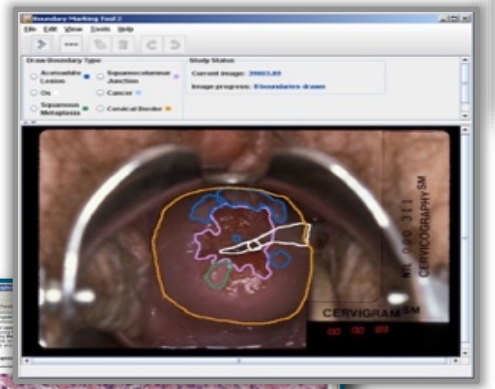


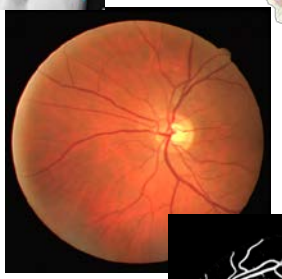


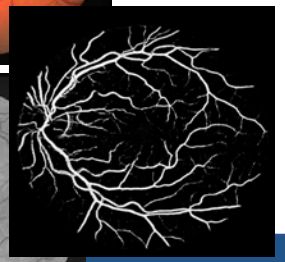


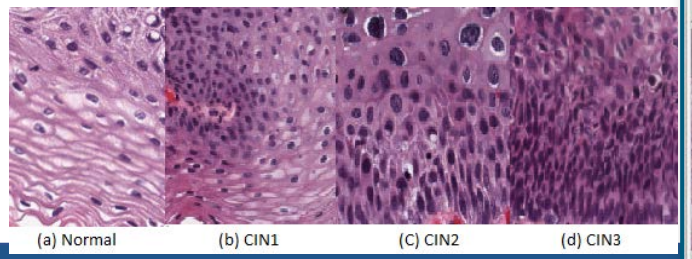












What do we want AI to be?

- Awesome!
- Transparent
  - As a user: be invisible
  - As a regulator (critic / cynic): not opaque about underlying technology
- Reproducible
- Explainable
- Improve existing processes

sameer.antani@nih.gov

Considerations:

- **Data**
  - Acquisition QA
  - Acquisition sensor sensitivity
  - Variety (Prevalence, coinfections)
  - Set Size, image (or datum) resolution
  - Sparsity
  - Relevance – How it answer the question? General AI or Narrow AI?
  - Modality (Data types, temporality)
  - Representing reality? (a perfect, curated and clean data set vs. dirty, real, missing)

sameer.antani@nih.gov

Considerations:

- **Algorithm Design**
  - AI Goals (Detection, Staging/Prognosis, Alerting)
  - Clinical purpose (clinician's aid, human-expert replacement, non-expert guidance)

- **Evaluation & Validation**
  - Evaluation measures – accuracy, recall, precision vs. sensitivity, specificity, and area under the operating curve
  - Role of statistics (sampling vs. exhaustive testing)
  - Training and test data partitioning, and collusion
  - Visualizing learning (internal layers vs. classification stage(s))
  - As the AI systems become more complex, how easily can these be teased out (particularly if the AI is building on prior work.)?

sameer.antani@nih.gov

Considerations:

- **Reproducibility**
  - How much non-private data is necessary to reflect AI performance?
  - Transparency in model design, data used, and training and test parameters

- **Development and Performance Consistency**
  - Dependency on GPU hardware,
  - Development platform,
  - Software version (libraries),
  - Data augmentation techniques

sameer.antani@nih.gov

Considerations:

- **Impact on Clinical Workflow**
  - Process change necessary?
  - Communicating with other (non-AI / AI) systems; standards
- **End-User Training**
  - User Complacency: “I took the action because the AI said so …”
  - Value of visualizations
- Continuous learning in AI (feedback), controls

sameer.antani@nih.gov

Questions? Thoughts? Comments?

sameer.antani@nih.gov